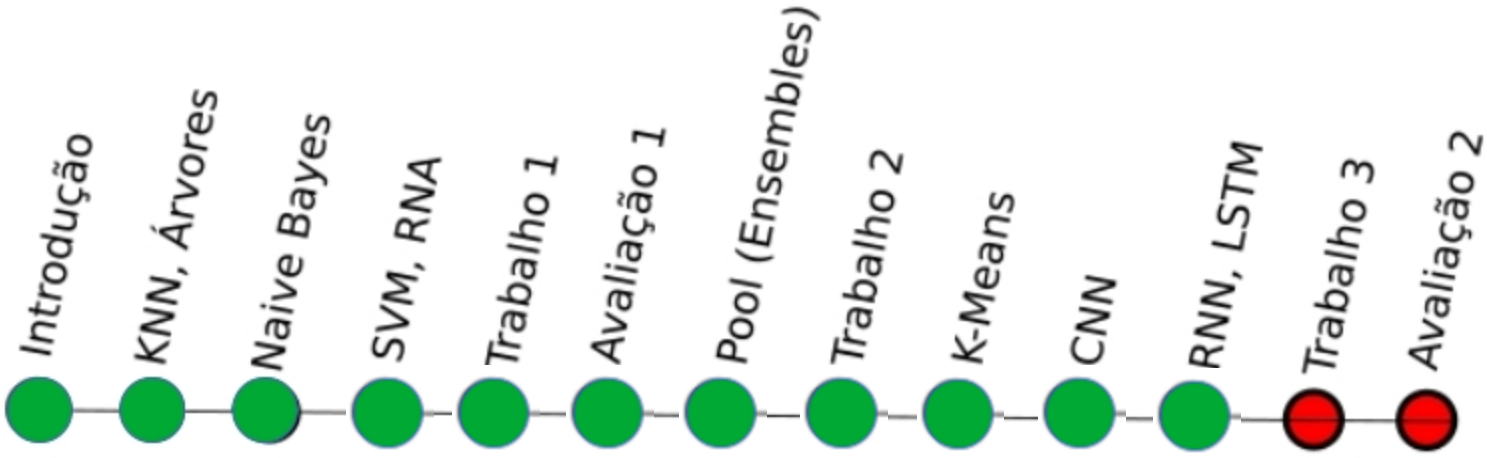
Tópico 04 – Deep Learning – Redes Recorrentes

Prof. André Gustavo Hochuli

gustavo.hochuli@pucpr.br
aghochuli@ppgia.pucpr.br

Tópicos

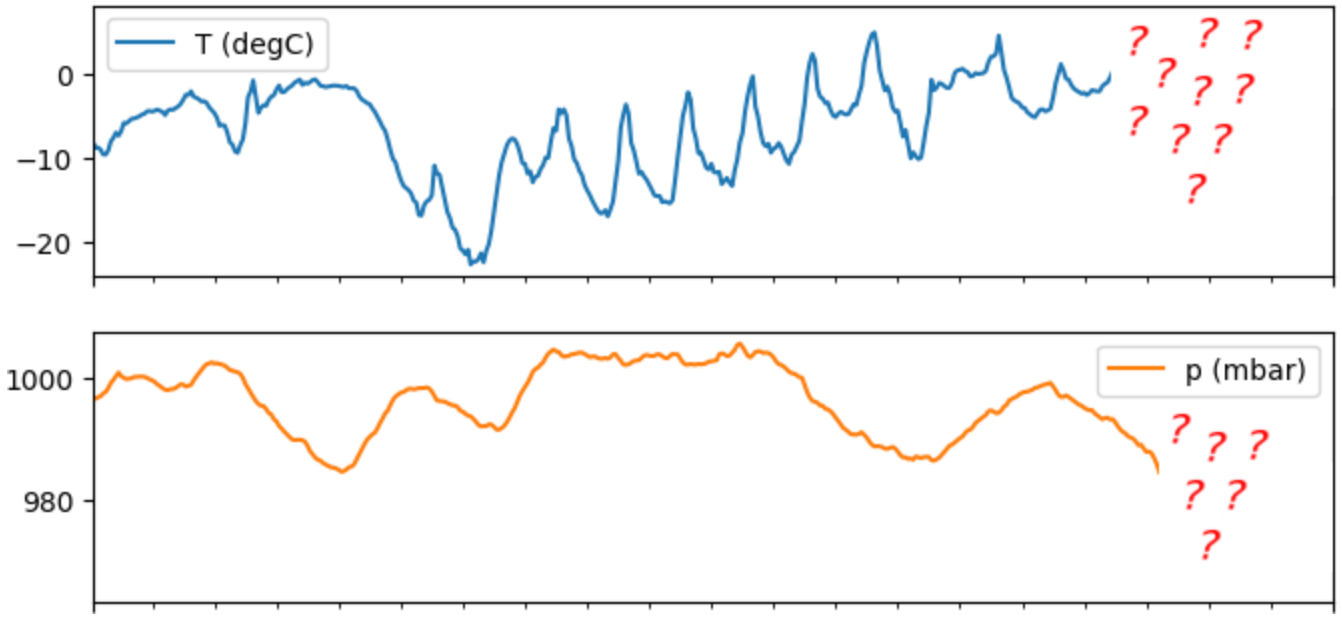
- Discussão Inicial
- Redes Recorrentes
  - RNN
  - LSTM
- Exercício



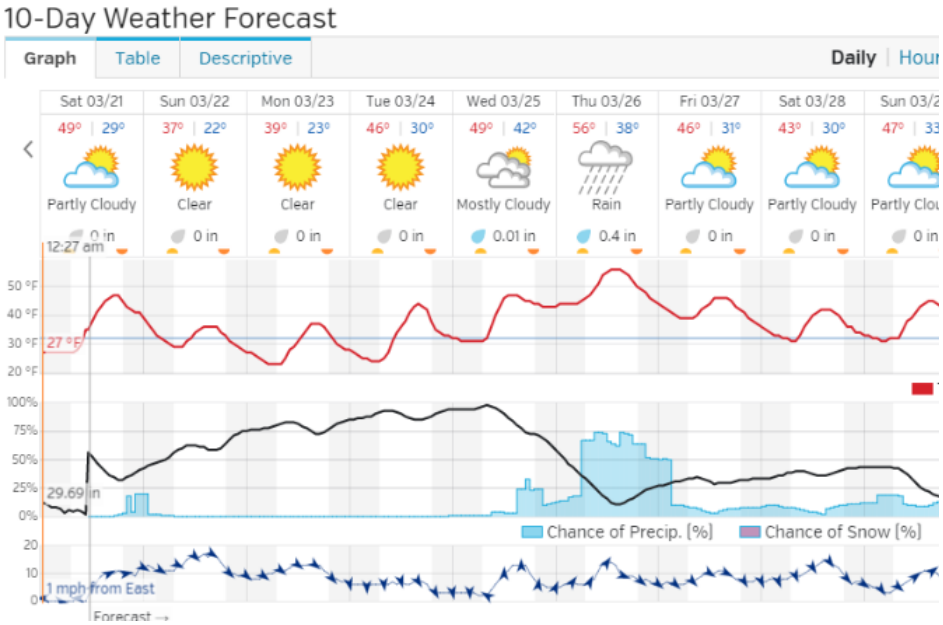
Discussão Inicial

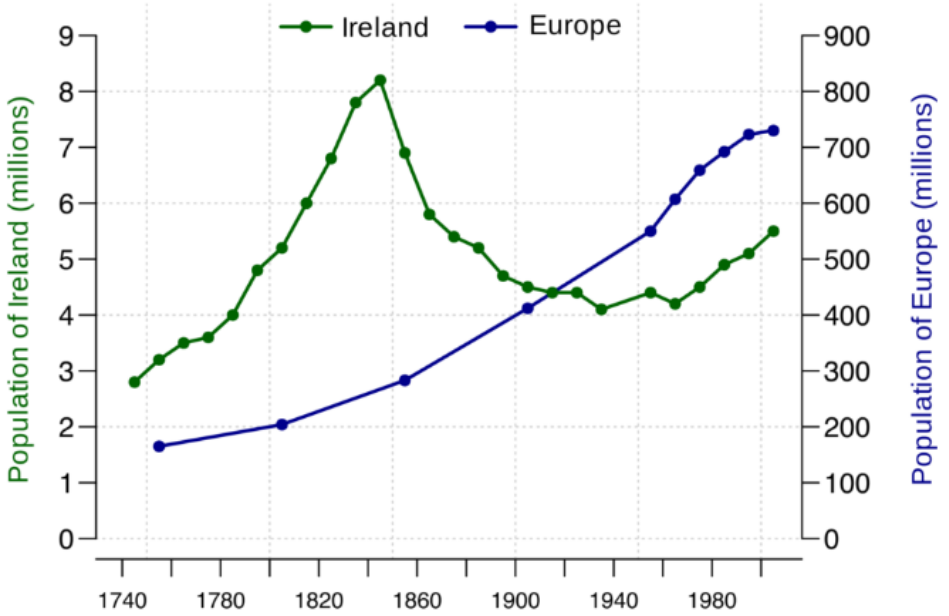
- E quando temos uma informação temporal ?
  - Ex: previsão do tempo, variação de preços, etc ?

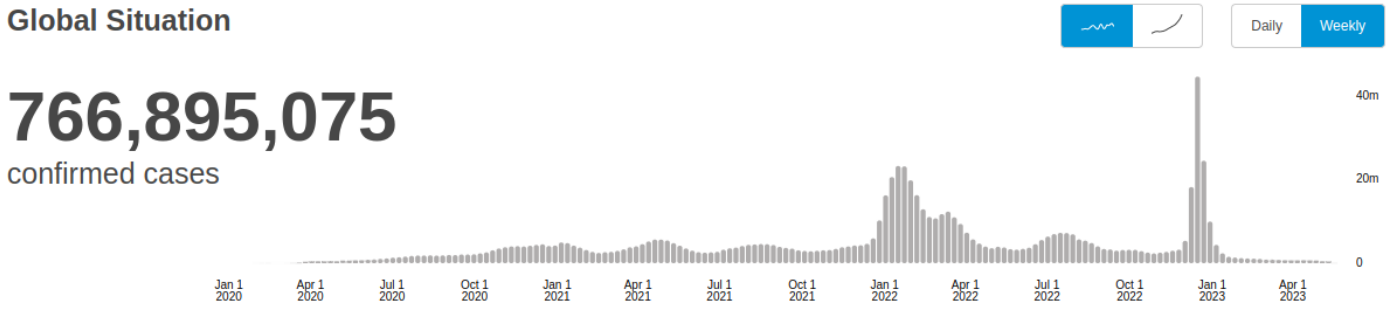
- Em certas aplicações, uma dependência temporal afeta a classe da instância



Discussão Inicial



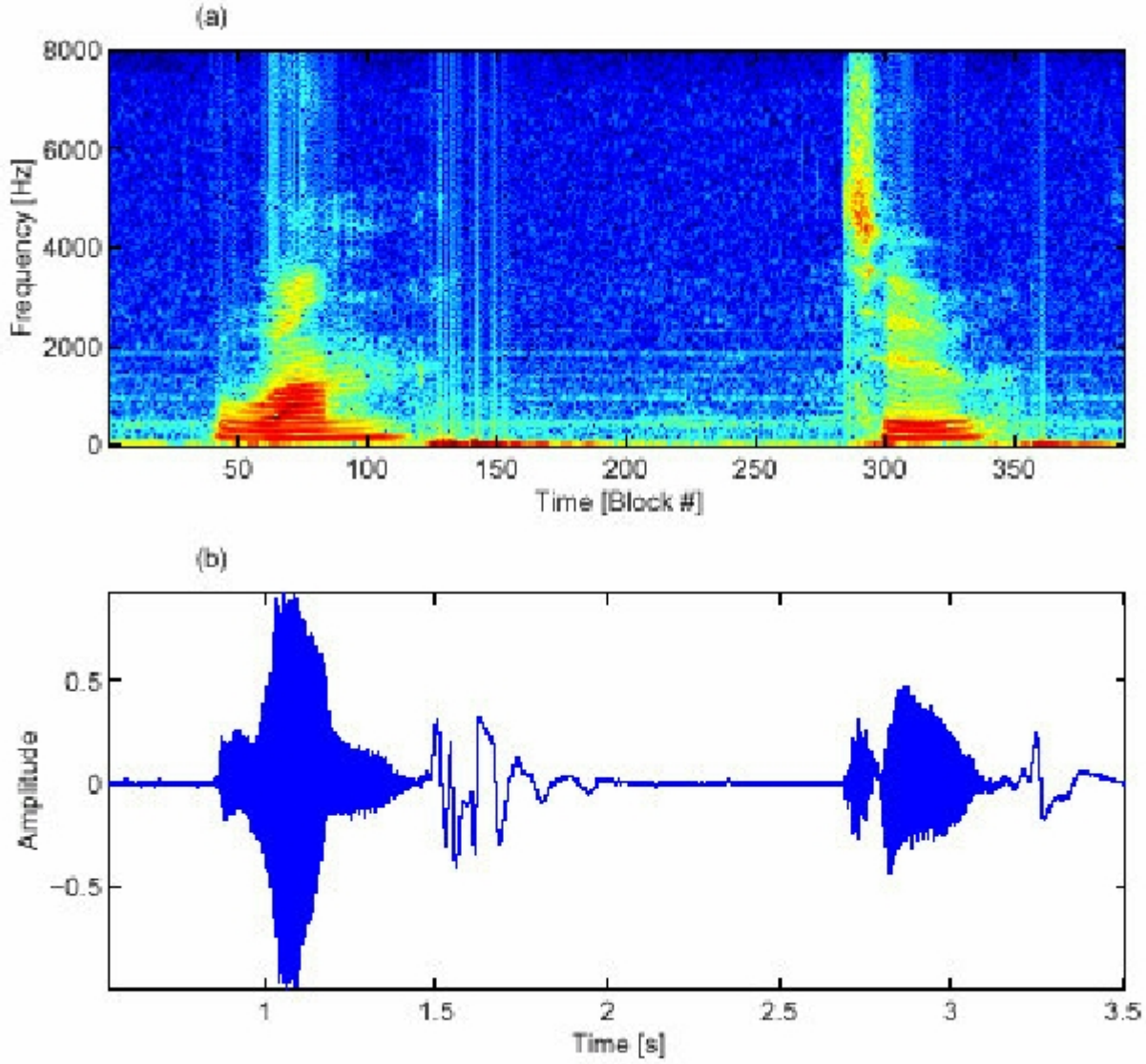




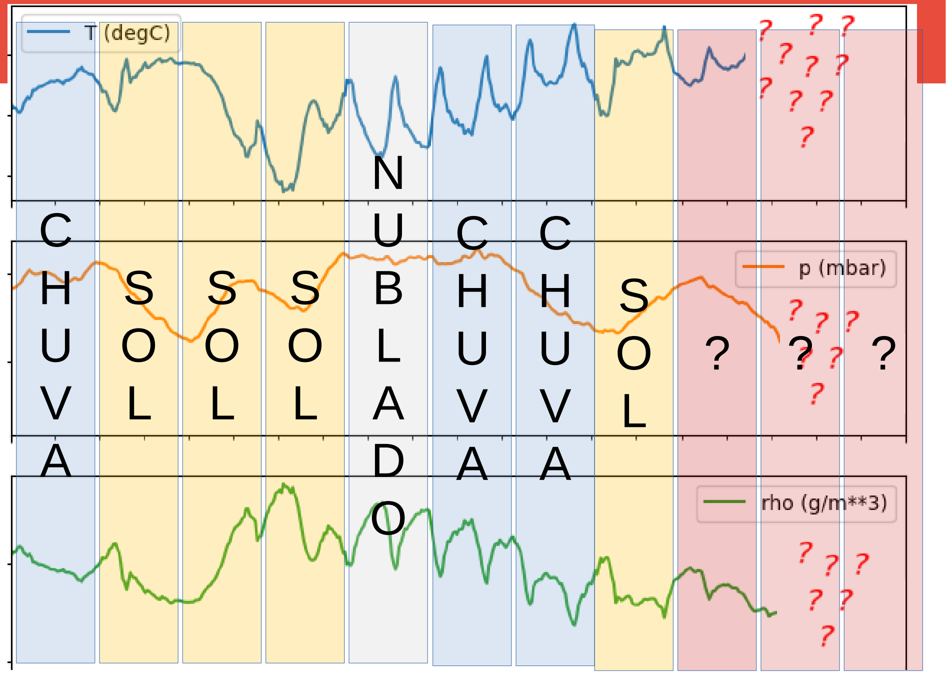
Aprendizado de Máquina - Prof. André Hochuli

Redes Recorrentes

Discussão Inicial



Discussão Inicial



- Classificação “Estática”
  - Instâncias são interpretadas isoladamente

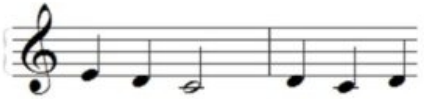
- Classificação Recorrente
  - A classe da instância anterior é importante para interpretação da instância atual

  - Exemplos
    - Imagem vs Vídeo

Redes Neurais Recorrentes

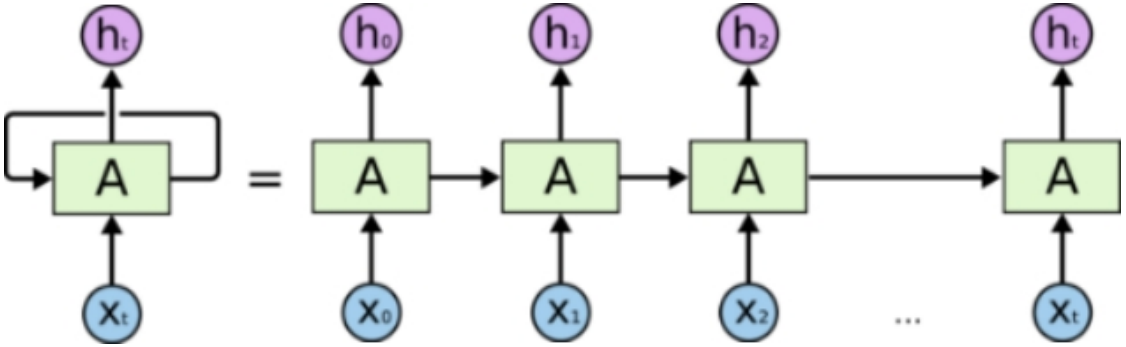
Redes Neurais Recorrentes

- Modelos capazes de interpretar sequência de dados
- Um conjunto de eventos determina a classe

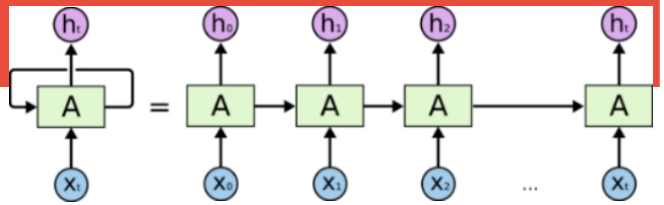
	x - input		y - output
Speech recognition		⟹	"Fuzzy Wuzzy was a bear. Fuzzy Wuzzy had no hair."
Music generation	∅	⟹	
Sentiment classification	"Decent effort. The plot could have been better."	⟹	
DNA sequence analysis	ACTGTACCCATGTGACTGCCC	⟹	ACTGTACCCATGTGACTGCCC
Machine translation	"El que no arriesga, no gana."	⟹	"If you don't take risks, you cannot win."
Video activity recognition		⟹	Running
Name entity recognition	"Ygritte says Jon Snow knows nothing."	⟹	"Ygritte says Jon Snow knows nothing."

Redes Neurais Recorrentes

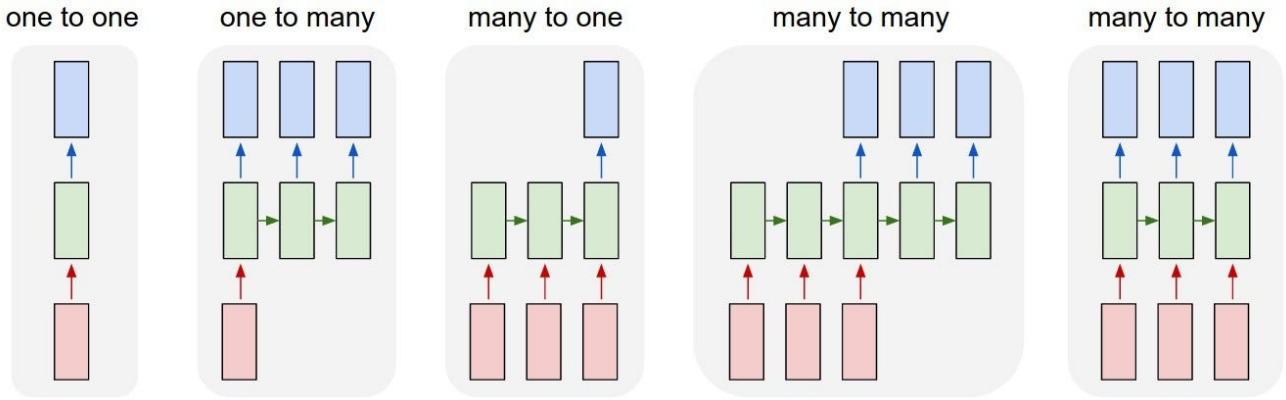
- Modelos ‘estáticos’ não codificam adequadamente a informação contextual de instâncias anteriores (série temporal)
- Solução: Recurrent Neural Networks
  - RNA propagando pesos + atributos
  - O estado H_t é produzido com base em $X_t + W_{t-1}$
  - Aprende a relação entre as instâncias e classes

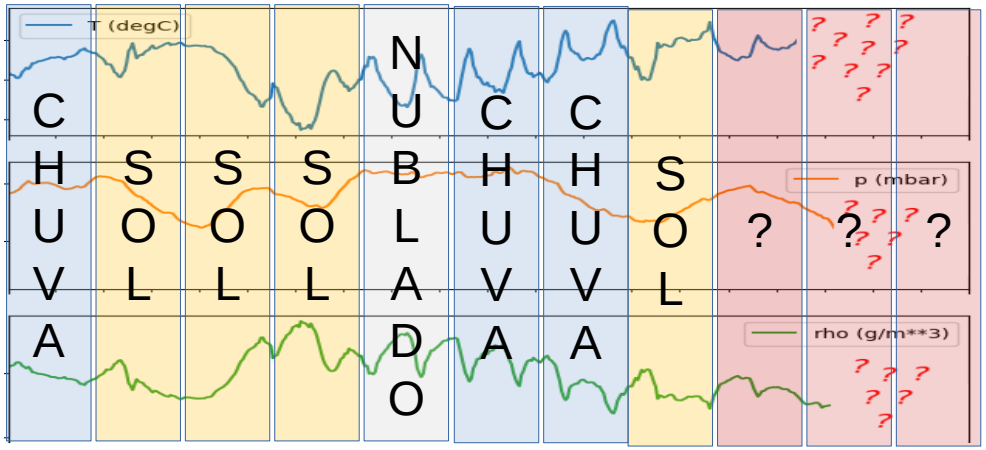


Redes Neurais Recorrentes

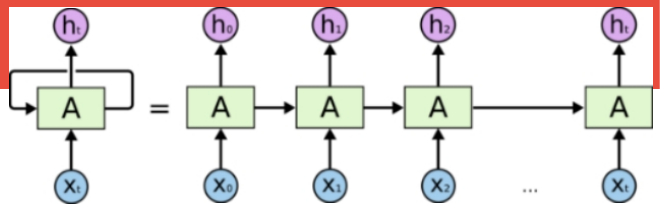


Abordagens:

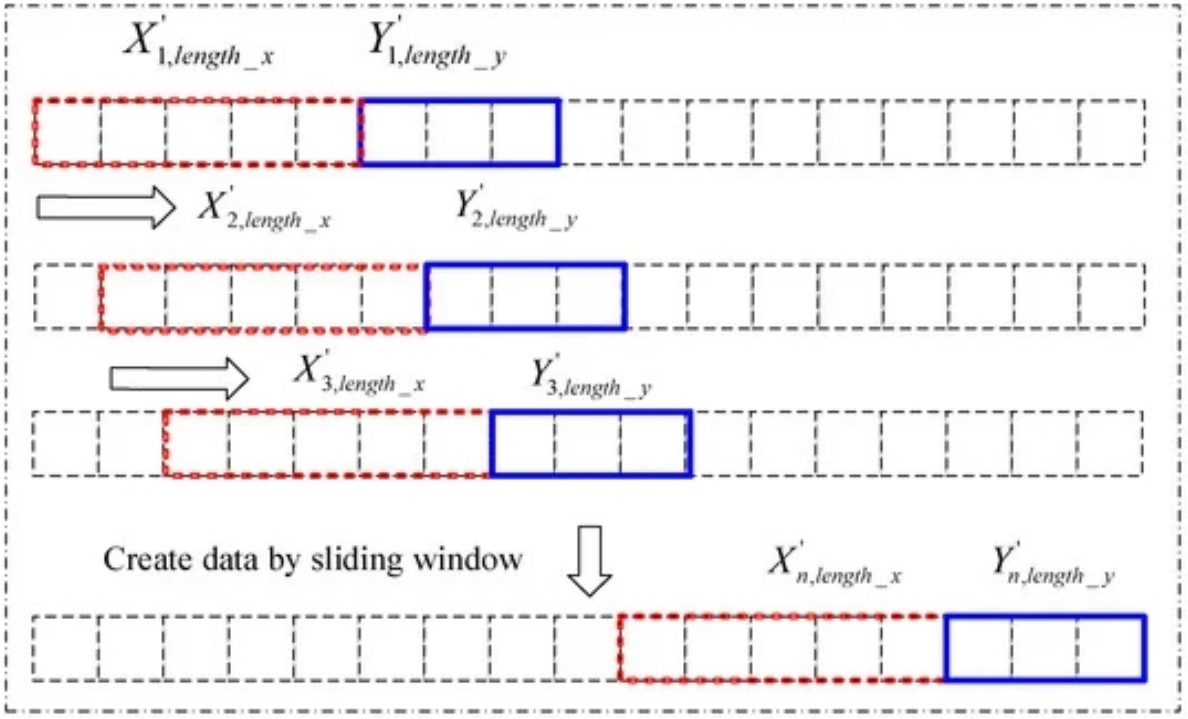




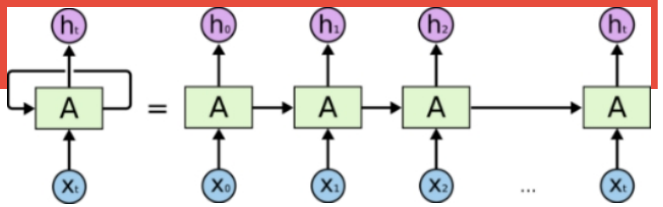
Redes Neurais Recorrentes



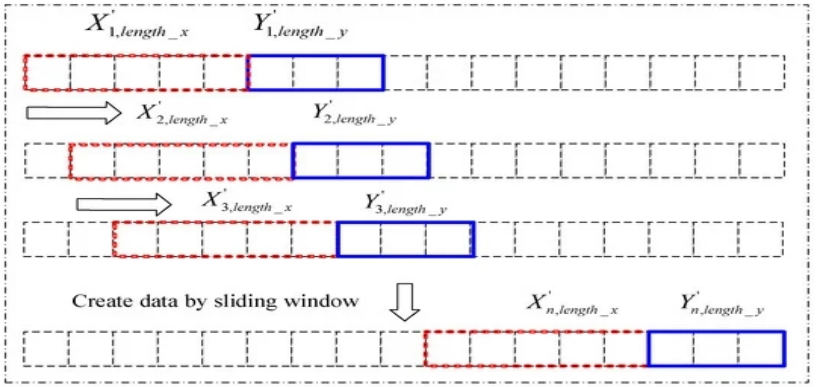
- Parâmetro: Tamanho da Janela
  - Determina a relação entre observação e predição

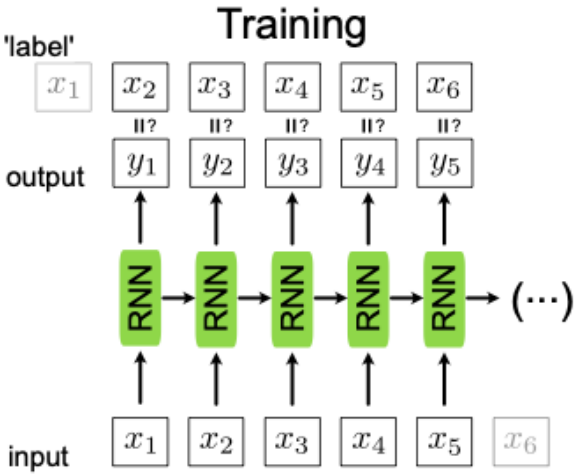


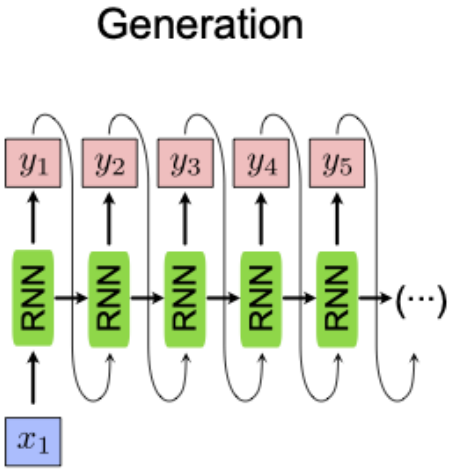
Redes Neurais Recorrentes

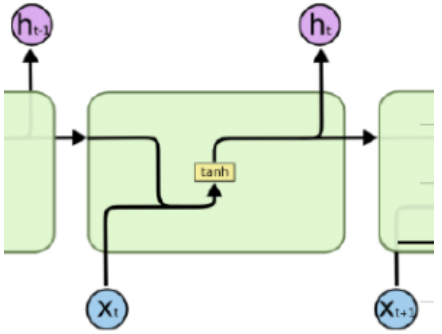


- Treinamento / Teste
  - Base de amostras são geradas deslizando a janela

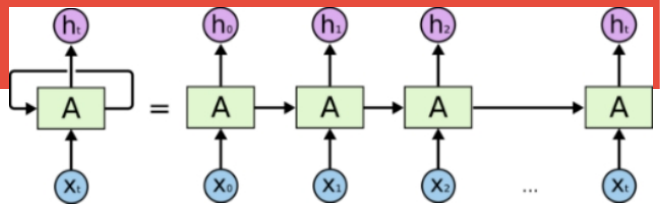




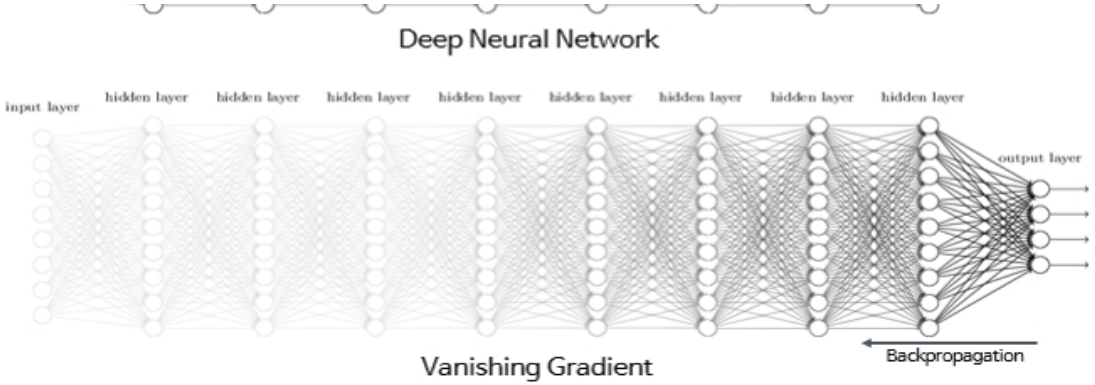


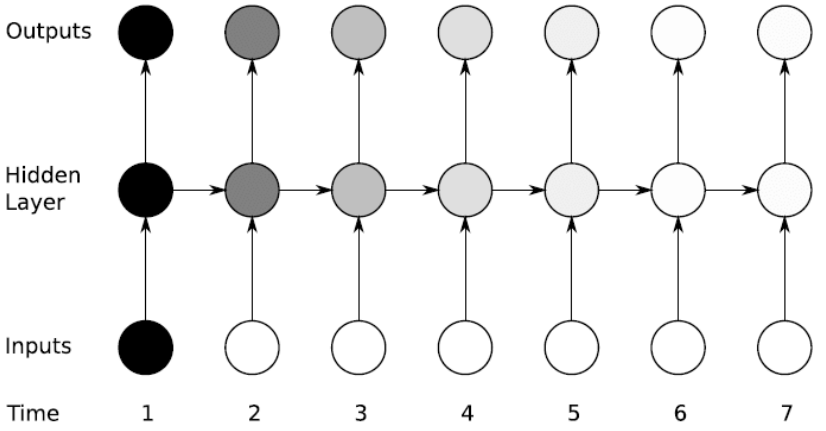


Redes Neurais Recorrentes

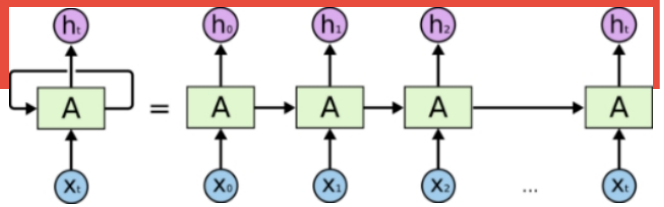


- Problema – Vanish Gradient
  - Em sequências grandes, o gradiente desaparece
  - O peso da informação ‘antiga’ decresce ao longo do tempo

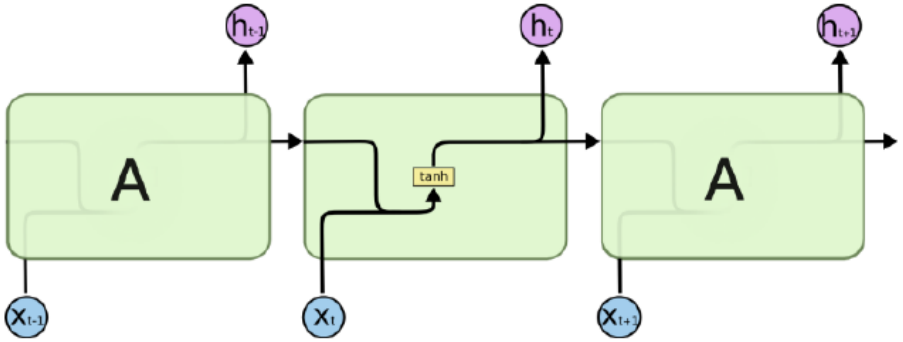


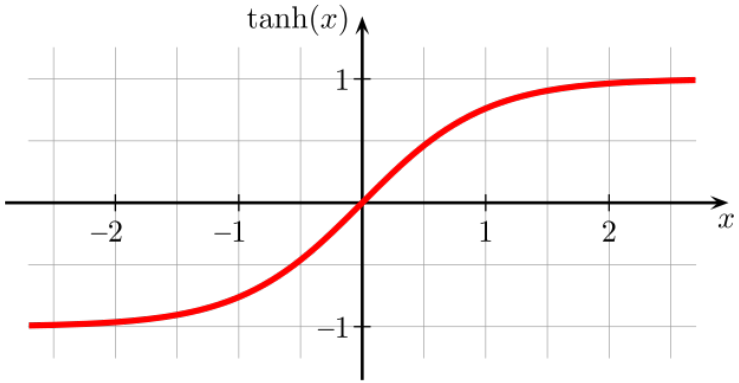


Redes Neurais Recorrentes



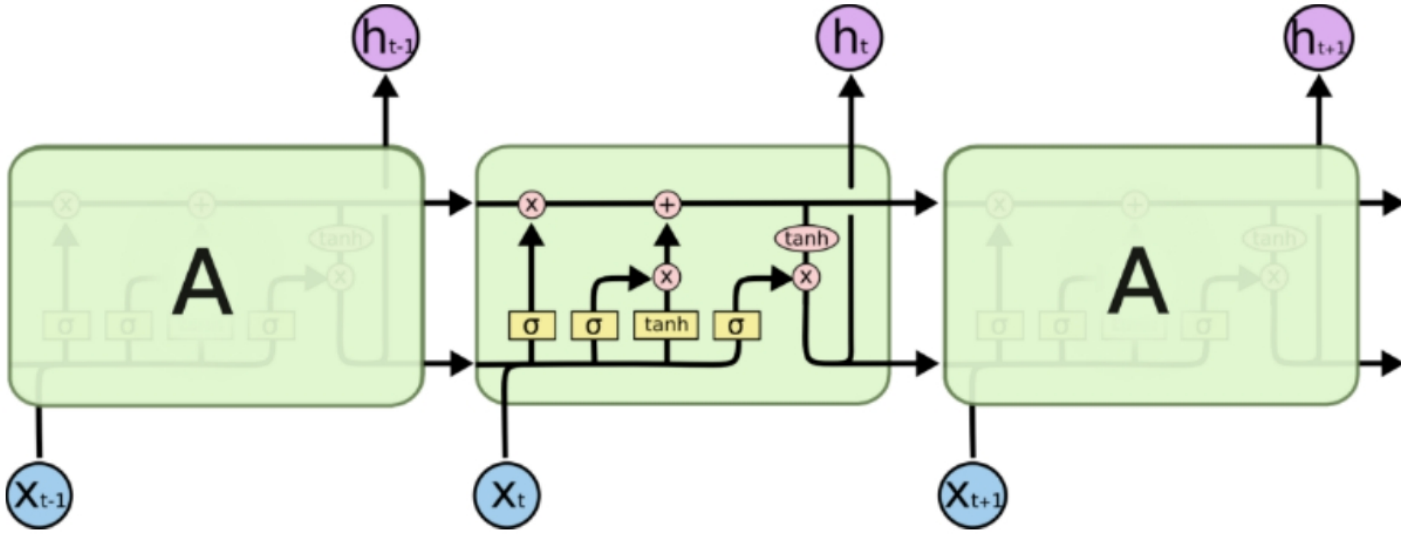
- Problema – Vanish Gradient
  - Tanh é uma função lenta



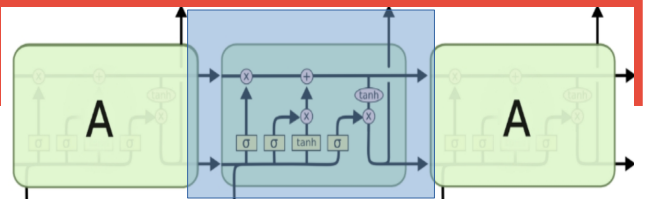


Redes Neurais Recorrentes

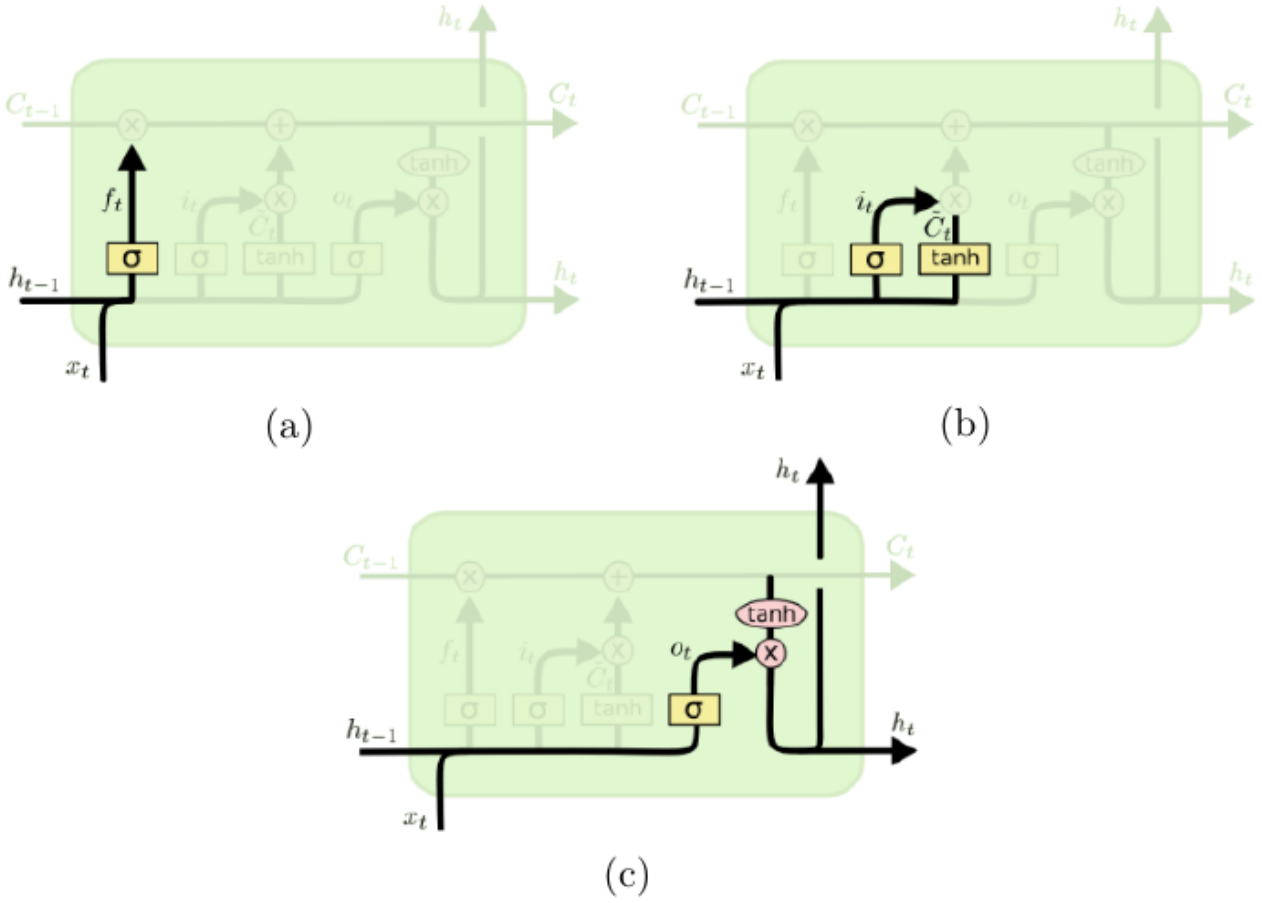
- Long-Short-Term-Memory (LSTM)
  - Fluxo Superior – Memória Longa
    - Multiplicação e **Soma** dos Pesos
  - Fluxo Inferior – Memória Curta
    - Multiplicação dos Pesos



Redes Neurais Recorrentes



- Gatilhos
  - Esquecimento (*)
  - Entrada (+)
  - Saída (*)



ura 2.14: Gatilhos LSTM: (a) Esquecimento (f_t), (b) Entrada (i_t) e (C) Saída (o_t).

Redes Neurais Recorrentes

- Let’s Code!
- [LINK]